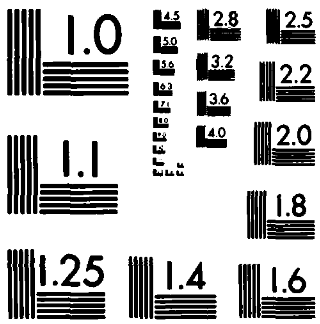
AD-A145 060 TESTING WHETHER NEW IS BETTER THAN USED OF A SPECIFIED AGE WITH RANDOMLY..(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS M HOLLANDER ET AL. DEC 83 1/1

UNCLASSIFIED FSU-STATISTICS-M669 AFOSR-TR-84-0716 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 84-0716

(3)

AD-A145 060

Testing Whether New is Better Than Used of a Specified Age, With Randomly Censored Data

by

Myles Hollander, Dong Ho Park, Frank Proschan

Florida State University, University of Nebraska, Florida State University

FSU Statistics Report M669

AFOSR Technical Report No. 83-164

December, 1983

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

F49620-82-K-0007

Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant Number ~~AFOSR-82-K-0007.~~ The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Key words and phrases: New better than used of a specified age, hypothesis test, randomly censored data, loss in efficiency due to censoring.

AMS 1980 *subject classifications:* Primary 62N05; secondary 62G10

84 08 30 030

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE:

3. DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S): FSU Statistics Report No. M669

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 84-0716

6a. NAME OF PERFORMING ORGANIZATION: Florida State University

6b. OFFICE SYMBOL *(If applicable)*:

7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research

6c. ADDRESS *(City, State, and ZIP Code)*: Dept of Statistics, Tallahassee FL 32306

7b. ADDRESS *(City, State, and ZIP Code)*: Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332

8a. NAME OF FUNDING / SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL *(If applicable)*: NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-82-K-0007

8c. ADDRESS *(City, State, and ZIP Code)*: Bolling AFB DC 20332

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A5	

11. TITLE *(Include Security Classification)*: TESTING WHETHER NEW IS BETTER THAN USED OF A SPECIFIED AGE, WITH RANDOMLY CENSORED DATA

12. PERSONAL AUTHOR(S): Myles Hollander, Dong Ho Park, and Frank Proschan

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Technical	FROM ______ TO ______	December 1983	16

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*: New better than used of a specified age; hypothesis test; randomly censored data; loss in efficiency due to censoring.

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Using randomly censored data, we develop a test of the null hypothesis that a new item has stochastically the same residual lifelength as does a used item of specified age t_0, versus the alternative hypothesis that a new item has stochastically greater residual lifelength than does a used item of age t_0. We also compare our test with a related test, developed for a complete-data model, in order to study the loss in efficiency because of censoring.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: CPT Brian W. Woodruff

22b. TELEPHONE *(Include Area Code)*: (202) 767- 5027

22c. OFFICE SYMBOL: NM

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE: UNCLASSIFIED

84 08 30 030

ABSTRACT

Using randomly censored data, we develop a test of the null hypothesis that a new item has stochastically the same residual lifelength as does a used item of specified age t_0, versus the alternative hypothesis that a new item has stochastically greater residual lifelength than does a used item of age t_0. We also compare our test with a related test, developed for a complete-data model, in order to study the loss in efficiency because of censoring.

1. INTRODUCTION

Suppose a cancer specialist believes that a patient newly diagnosed as having a certain type of cancer has a smaller chance of survival than does a patient who has survived 5 years following a similar initial diagnosis. How can such a claim be tested? To address such a testing problem, Hollander, Park, and Proschan (1983) introduced the class G_1 of "new better than used at t_0" life distributions and the dual class G_2 of "new worse than used at t_0" life distributions.

G_1, *the new better than used at* t_0 *class.* Let $t_0>0$. A life distribution F (i.e., a distribution such that $F(x) = 0$ for all $x \le 0$) is new better than used at t_0 if

$$\bar{F}(x+t_0) \le \bar{F}(x)\bar{F}(t_0) \quad \text{for all} \quad x \ge 0, \tag{1.1}$$

where $\bar{F} = 1-F$ denotes the survival function.

G_2, *the new worse than used at* t_0 *class.* A life distribution F is new worse than used at t_0 if the first inequality in (1.1) is reversed.

We now describe another context in which the concept of new worse than used at t_0 finds practical application. For many electronic components, experience shows that the failure rate of a component is higher during the so-called "infant-mortality phase" (say $[0,t_0]$) than after t_0. This could be the consequence of the fact that there are really *two* populations which have been mixed during manufacture - one population consists of well made units, while the second population consists of poorly made units whose defects show up immediately upon initial use or shortly thereafter.

The practical procedure used by manufacturers of such components is to "burn-in" a portion of their product during $[0, t_0]$; surviving components of age t_0 sell at a *higher* price than new untested components. The purchaser who needs higher than usual quality for say, space vehicle assembly, safety equipment, surgical devices, etc., is willing to pay for the statistically higher quality units of age t_0.

The manufacturer or one of his high volume customers may be highly motivated to carry out the statistical test proposed in this paper to corroborate his engineering experience and judgement.

$G_1[G_2]$ is related to but contains and is much larger than the class $H_1[H_2]$ of "new better than used" ["new worse than used"] distributions defined below.

H_1, *the new better than used class.* A life distribution F is new better than used if

$$\bar{F}(x+t) \le \bar{F}(x)\bar{F}(t) \quad \text{for all} \quad x,\ t \ge 0. \tag{1.2}$$

H_2, *the new worse than used class.* A life distribution F is new worse than used if it satisfies (1.2) with the first inequality reversed.

Thus the G_1 property states that a used item of age t_0 has stochastically smaller residual lifelength than does a new item whereas the H_1 property states that a used item of any age has stochastically smaller residual life-length than does a new item. Analogous interpretations hold for G_2 and H_2.

The only members in $H_1 \cap H_2$ are the exponential distributions. Theorem 2 of Marsaglia and Tubilla (1975) shows that the only members of $G_1 \cap G_2$ are (a) exponential distributions, (b) survival functions $\bar{F}$ for which $\bar{F}(0)=1$ and $\bar{F}(t_0)=0$, and (c) survival functions of the form $\bar{F}^*(x)=\bar{F}(x)$ for $0 \le x < t_0$, $= \bar{F}^j(t_0)\bar{F}(x-jt_0)$ for $jt_0 \le x < (j+1)t_0$, $j=0,1,2,\ldots$, where $\bar{F}$ is a survival function defined for $x \ge 0$. If F has a density function on $[0,t_0]$, then the failure rate of F^* is periodic with period t_0.

Some examples of new better than used life distributions are the Weibull where $F_\theta(t) = 1-\exp(-(\lambda t)^\theta)$, $t \ge 0$, $\lambda \ge 0$, $\theta \ge 1$ and the gamma where $F_\theta(t) = \int_0^t \lambda^\theta x^{\theta-1} \exp(-\lambda x)/\Gamma(\theta)$, $t \ge 0$, $\lambda \ge 0$, $\theta \ge 1$.

Let C* be the class of life distributions which are not new better than used but are new better than used at t_0. Hollander, Park, and Proschan (1983) give the following method of constructing some distribution functions in C*. Suppose that G is a new better than used distribution with failure rate function $r_G(x) > 0$ for $0 \le x < \infty$. Let F have a failure rate function r_F satisfying (i)$r_F(x) \le r_G(x)$ for $0 \le x \le t_0$, (ii)$r_F(x) = r_G(x)$ for $t_0 \le x < \infty$, and (iii)$r_F(x)$ is strictly decreasing for $0 \le t \le t_1$, where $0 < t_1 < t_0$. Then $F \in C^*$. To develop an example of this construction, let $r_G(x) = 1$ for $0 \le x < \infty$, and let $r_F(x) = 1-(\theta/t_0)x$ for $0 \le x < t_0$ and $0 < \theta \le 1$, and $r_F(x) = 1$ for $t_0 \le x < \infty$. (We do not let θ exceed 1 since we want to ensure that $r_F(x)$ remains positive as $x \to t_0$.) Then r_F satisfies (i),(ii), and (iii) and thus F is in C*. The corresponding survival function can be written as

$$\begin{aligned}\bar{F}_\theta(t) &= \exp[-\{t-\theta(2t_0)^{-1}t^2\}],\ 0 < \theta \le 1,\ 0 \le t < t_0, \\ &= \exp[-\{t-\theta(2)^{-1}t_0\}],\ 0 < \theta \le 1,\ t \ge t_0.\end{aligned}$$

Hollander, Park, and Proschan (1983) derived a test of

$$H_0:\ F \in G_1 \cap G_2 \tag{1.3}$$

versus

$$H_A:\ F \in G_1^* \tag{1.4}$$

where

$$G_1^* = \{F:\ \bar{F}(x+t_0) \le \bar{F}(x)\bar{F}(t_0) \text{ for all } x \ge 0 \text{ and inequality holds for some } x \ge 0\}.$$

The null hypothesis H_0 asserts that a new item has stochastically the same residual lifelength as does a used item of age t_0. (Equivalently F satisfies H_0 if it satisfies (1.1) where the first inequality is replaced by an equality.) The alternative hypothesis H_A asserts that a new item has stochastically greater residual lifelength than does a used item of age t_0. The test proposed by Hollander, Park, and Proschan was for the model where we obtain a complete random sample from the life distribution F. In the present paper we propose a test of H_0 versus H_A based on a randomly right-censored sample. The test is derived in Section 2. In Section 3 we compare the tests for the uncensored and censored models and obtain a measure of the loss in efficiency incurred because of censoring. Section 4 contains an example.

2. A TEST OF H_0 VERSUS H_A USING INCOMPLETE DATA

Let $X_1, \ldots, X_n$ be independent and identically distributed according to a continuous life distribution F, and $Y_1, \ldots, Y_n$ be independent and identically distributed according to a continuous censoring distribution H. Also, let the X's be independent of the Y's. The censoring distribution H is typically, though not necessarily, unknown and is treated as a nuisance parameter. In the randomly censored model we do not observe $X_1, \ldots, X_n$, but instead, we observe the pairs (Z_i, δ_i), i = 1, ..., n, where

$$Z_i = \min(X_i, Y_i)$$

and

$$\delta_i = \begin{cases} 1 & \text{if } Z_i = X_i \\ 0 & \text{if } Z_i = Y_i . \end{cases}$$

Our test is based on an estimator of the parameter

$$T(F) = \int_0^\infty \{\bar{F}(x+t_0) - \bar{F}(x)\bar{F}(t_0)\}dF(x)$$

$$= \int_0^\infty \bar{F}(x+t_0)dF(x) - \frac{1}{2}\bar{F}(t_0) .$$

Under H_0, $T(F) = 0$ whereas under H_A, $T(F) \leq 0$. In fact, since F is assumed continuous, under H_A, $T(F)$ is strictly less than 0. $T(F)$ gives a measure of the deviation of F from H_0 towards H_A. It is natural to base a test of H_0 versus H_A on a consistent estimator of $T(F)$, and we utilize $T(\hat{F}_n)$ where $\hat{F}_n$ is the Kaplan-Meier (1958) estimator. (Such an approach has been used by many authors including Efron (1967) in the context of a two-sample location test, by Koziol and Green (1976) and Csörgő and Horvath (1981) in testing goodness-of-fit, and by Chen, Hollander, and Langberg (1983a) in testing whether new is better than used.)

Under the assumption that F is continuous, the Kaplan-Meier estimator can be expressed as

$$\hat{\bar{F}}_n(t) = \prod_{\{i:Z_{(i)} \leq t\}} \{(n-i)(n-i+1)^{-1}\}^{\delta_{(i)}}, \quad t \in [0, Z_{(n)}),$$

where $Z_{(0)} \equiv 0 < Z_{(1)} < \dots < Z_{(n)}$ denote the ordered Z's and $\delta_{(i)}$ is the δ corresponding to $Z_{(i)}$. Here we treat $Z_{(n)}$ as an uncensored observation, whether it is uncensored or censored. When censored observations are tied with uncensored observations, our convention for ordering the Z's is to treat the uncensored observations as preceding the censored observations.

Weak convergence of the Kaplan-Meier estimator, regarded as a stochastic process, has been established by Efron (1967), Breslow and Crowley (1974), Meier (1975), and Gill (1983). Strong consistency of the estimator was proved by Peterson (1977) and Langberg, Proschan, and Quinzi (1981). Exact small-sample moments of $\hat{\bar{F}}_n$, under a model of proportional hazards, were obtained by Chen, Hollander, and Langberg (1982).

Our test statistic is

$$T_n^c \overset{\text{def.}}{=} T(\hat{F}_n) = \int_0^\infty \hat{\bar{F}}_n(x+t_0)\,d\hat{F}_n(x) - \tfrac{1}{2}\hat{\bar{F}}_n(t_0).$$

For computational purposes we may rewrite T_n^c as

$$T_n^c = \sum_{i=1}^{n} \left\{ \left[\prod_{\{k:Z_{(k)} \le Z_{(i)}+t_0\}} \{(n-k)(n-k+1)^{-1}\}^{\delta(k)} \right] \right.$$

$$\left. \cdot \left[\prod_{r=1}^{i-1} \{(n-r)(n-r+1)^{-1}\}^{\delta(r)} \right] \cdot \left[1-\{(n-i)(n-i+1)^{-1}\}^{\delta(i)} \right] \right\}$$

$$- \frac{1}{2} \prod_{\{k:Z_{(k)} \le t_0\}} \{(n-k)(n-k+1)^{-1}\}^{\delta(k)}.$$

Asymptotic normality of $n^{1/2}\{T_n^c - T(c)\}$ can be established under the assumptions

(A.1) The support of both F and H is $[0,\infty)$.

(A.2) $\sup\{[\bar{F}(x)]^{1-\varepsilon}[\bar{H}(x)]^{-1},\ x \in [0,\infty)\} < \infty$ for some $\varepsilon \ge 0$.

Condition (A.2) restricts the amount of censoring. For example, in the proportional hazards model where $\bar{H}(t) = \{\bar{F}(t)\}^{\beta}$, (A.2) implies that $\beta < 1$, which means that the expected proportion of censored observations $\beta/\beta+1$ must be less than .5.

We now state the main result of this section and then sketch its proof.

THEOREM 1. Assume that (A.1) and (A.2) hold. Then $n^{1/2}\{T_n^c - T(F)\}$ converges in distribution to a normal random variable with mean 0 and variance σ_c^2 given by equation (2.1) below.

Sketch of Proof. For n = 1, 2, ..., the expression for $T_n^c - T(F)$ can be written as:

$$T_n^c - T(F) = \int\{\hat{\bar{F}}_n(x+t_0)-\bar{F}(x+t_0)\}d\hat{F}_n(x)-\int\{\hat{\bar{F}}_n(x+t_0)-\bar{F}(x+t_0)\}dF(x)$$

$$+ \int\bar{F}(x+t_0)d\hat{F}_n(x) - \int\bar{F}(x+t_0)dF(x)$$

$$+ \int\hat{\bar{F}}_n(x+t_0)dF(x) - \int\bar{F}(x+t_0)dF(x)$$

$$- \tfrac{1}{2}\hat{\bar{F}}_n(t_0) + \tfrac{1}{2}\bar{F}(t_0).$$

Unless otherwise specified, all integrals range over $(0, \infty)$. Upon integration by parts and change of variables, we have:

$$\int\bar{F}(x+t_0)d\hat{F}_n(x)-\int\bar{F}(x+t_0)dF(x) = -\int\hat{\bar{F}}_n(x-t_0)dF(x)+\int\bar{F}(x-t_0)dF(x),\ n = 1, 2, \ldots .$$

Thus for n = 1, 2, ...,

$$n^{1/2}(T_n^c-T(F)) = B_{n,1} + B_{n,2},$$

where

$$B_{n,1} = \int n^{1/2}\{\hat{\bar{F}}_n(x+t_0)-\bar{F}(x+t_0)\}d\hat{F}_n(x)-\int n^{1/2}\{\hat{\bar{F}}_n(x+t_0)-\bar{F}(x+t_0)\}dF(x),$$

and

$$B_{n,2} = \int n^{1/2}[\hat{\bar{F}}_n(x+t_0)-\bar{F}(x+t_0)-\{\hat{\bar{F}}_n(x-t_0)-\bar{F}(x-t_0)\}-\tfrac{1}{2}\{\hat{\bar{F}}_n(t_0)-\bar{F}(t_0)\}]dF(x).$$

Using Theorem 2.1 of Gill (1983) and arguments similar to those in the proofs of Lemmas 2.2 and 2.3 of Chen, Hollander, and Langberg (1983a), we can show that

(a) $B_{n,1}$ converges in probability to zero,

(b) $B_{n,2}$ converges in distribution to the random variable

$Z_0 = \int\{\phi(x+t_0)-\phi(x-t_0) - \frac{1}{2}\phi(t_0)\}dF(x)$, and

(c) The random variable Z_0 is normal with mean 0 and finite variance

$$\sigma_c^2 = \iint E\{\phi(x+t_0)-\phi(x-t_0) - \tfrac{1}{2}\phi(t_0)\} \cdot \{\phi(u+t_0)-\phi(u-t_0) - \tfrac{1}{2}\phi(t_0)\}dF(x)dF(u). \quad (2.1)$$

Here $\{\phi(t),\ t\in(0,\infty)\}$ is the Gaussian process which is the limit of the Kaplan-Meier estimator regarded as a stochastic process. The mean of $\phi(t)$ is zero and its covariance kernel is

$$E\phi(t)\phi(s) = \bar{F}(t)\bar{F}(s) \int_0^s \{\bar{K}(z)\bar{F}(z)\}^{-1}dF(z),\ 0 \le s < t < \infty, \quad (2.2)$$

where $\bar{K}(t) = \bar{F}(t)\bar{H}(t)$. This concludes the sketch.

The null asynptotic mean of $n^{1/2}T_n^c$ is zero, independent of the distributions F and H. However, the null asymptotic variance of $n^{1/2}T_n^c$ depends on both F and H and thus it must be estimated from the incomplete observations $(Z_1,\delta_1), \ldots, (Z_n,\delta_n)$. Under H_0 it can be shown, after straightforward but tedious integration using the expressions for σ_c^2 and $E\phi(t)\phi(s)$ given in (2.1) and (2.2), that the null asymptotic variance of $n^{1/2}T_n^c$ is:

$$\begin{aligned}\sigma_{c0}^2 &= (1/4)\bar{F}^2(t_0)\int_0^\infty \bar{F}^3(z)\{\bar{K}(z+t_0)\}^{-1}dF(z) \\ &+ (1/4)\bar{F}^2(t_0)\int_0^\infty \bar{F}^3(z)\{\bar{K}(z)\}^{-1}dF(z) - (1/2)\bar{F}^4(t_0)\int_0^\infty \bar{F}^3(z)\{\bar{K}(z+t_0)\}^{-1}dF(z).\end{aligned} \quad (2.3)$$

If there is no censoring, that is, if $\bar{K}(z) = \bar{F}(z)$ for $z \in [0, \infty)$, then σ_{c0}^2 reduces to $(1/12)\bar{F}(t_0)+(1/12)\bar{F}^2(t_0)-(1/6)\bar{F}^3(t_0)$. This expression agrees with the null asymptotic variance σ_0^2 of the statistic $n^{1/2}T(\hat{G}_n)$ (where $\hat{G}_n$ is the empirical distribution function of a random sample from F) advanced by Hollander, Park, and Proschan (1983) for testing H_0 in the complete data case. Expression (2.3) can be simplified, by a change of variable in the first and third terms, to

$$\sigma_{c0}^2 = (1/4)\{\bar{F}(t_0)\}^{-2}\int_{t_0}^{\infty} \bar{F}^3(u)\{\bar{K}(u)\}^{-1}dF(u)$$

$$+ (1/4)\bar{F}^2(t_0)\int_0^{\infty}\bar{F}^3(u)\{\bar{K}(u)\}^{-1}dF(u) - (1/2)\int_{t_0}^{\infty} \bar{F}^3(u)\{\bar{K}(u)\}^{-1}dF(u). \tag{2.4}$$

To obtain our estimator of σ_{c0}^2 we introduce some notation. Let $Z_{(1)} \le \ldots \le Z_{(n)}$ denote the ordered Z-values. Let K_n denote the empirical distribution function of the Z-values. Thus $nK_n(t)$ = "number of Z-values $\le t$." Since $T_n^c = 0$ when $Z_{(n)} \le t_0$, we will assume our sample is such that $Z_{(n)} > t_0$. Replacing $\bar{F}$ by $\hat{\bar{F}}_n$, $\bar{K}$ by K_n, and ∞ by $Z_{(n)}$ in (2.4) yields the estimate $\hat{\sigma}_{cn}^2$ defined by (2.5).

$$\hat{\sigma}_{cn}^2 = [(1/4)\{\hat{\bar{F}}_n(t_0)\}^{-2} - (1/2)]$$

$$\cdot \sum_{\{i:t_0\le W_{(i)}\le Z_{(n)}\}} \hat{\bar{F}}_n^3(W_{(i)})\{\bar{K}_n(W_{(i)}-)\}^{-1}\{\hat{F}_n(W_{(i)})-\hat{F}_n(W_{(i-1)})\}$$

$$+ (1/4)\hat{\bar{F}}_n^2(t_0) \sum_{\{i:1\le i\le \tau(n)\}} \tag{2.5}$$

$$\cdot \hat{\bar{F}}_n^3(W_{(i)})\{\bar{K}_n(W_{(i)}-)\}^{-1}\{\hat{F}_n(W_{(i)})-\hat{F}_n(W_{(i-1)})\},$$

where $W_{(0)} \equiv 0 < W_{(1)} < W_{(2)} < \ldots < W_{(\tau(n))}$ are the ordered observed failure times, and $\tau(n) = \sum_{i=1}^{n} \delta_i$ is the total number of failures among the n observations.

We are unable to prove that $\hat{\sigma}_{cn}^2$ consistently estimates σ_{c0}^2 under our (A.1), (A.2) assumptions, but we have investigated properties of $\hat{\sigma}_{cn}^2$ via a limited Monte Carlo study. Table 1 investigates the accuracy of $\hat{\sigma}_{cn}^2$ as an estimator of σ_{c0}^2 and the accuracy of the normal approximation in the cases where F is exponential with scale parameter 1 and H is exponential with scale parameter λ, for the choices $\lambda=.1$ and $\lambda=1/3$, with $t_0=.6,1$, and $n=100, 150,200$. Column 2 of Table 1 gives the average value of $\hat{\sigma}^2_{cn}$, averaged over 1,000 Monte Carlo replications. Column 3 gives the sample standard deviation s of the 1,000 $\hat{\sigma}^2_{cn}$ values. It is seen that $\hat{\sigma}^2_{cn}$ tends to be below the true value σ^2_{c0}, but the estimator improves as n increases.

The approximate α-level test of H_0 versus H_A, which rejects H_0 in favor of H_A if $n^{1/2}T^c_n\hat{\sigma}^{-1}_{cn} \le -z_\alpha$ and accepts H_0 otherwise, is called the NBU-t_0 test. The approximate α-level test of H_0 versus the alternative that a new item has stochastically smaller residual lifelength than does a used item of age t_0 is called the NWU-t_0 test. The NWU-t_0 test rejects H_0 if $n^{1/2}T^c_n\hat{\sigma}^{-1}_{cn} \ge z_\alpha$ and accepts H_0 otherwise. Here z_α is the upper α-percentile point of a standard normal distribution. Columns 4,5,6,7,8,9 pertain to the convergence to asymptotic normality of the standardized test statistic $T^* = n^{1/2}T^c_n\hat{\sigma}^{-1}_{cn}$. Columns 4,5,6 give estimated probabilities of the events $\{T^* \le -z_\alpha\}$, and columns 7,8,9 give estimated probabilities of the events $\{T^* \ge z_\alpha\}$, $\alpha = .10,.05,.01$. It is seen that the convergence to asymptotic normality is slow. The probability α assigned to the event $\{T^* \le -z_\alpha\}$ by the normal approximation is less than the corresponding Monte Carlo estimate $\hat{P}\{T^* \le -z_\alpha\}$. Thus the NBU-$t_0$ test based on the normal approximation tends to give P values that are less than the true P values. The probability α assigned to the event $\{T^* \ge z_\alpha\}$ by the normal approximation is greater than the corresponding Monte Carlo estimate $\hat{P}\{T^* \ge z_\alpha\}$. Thus the NWU-$t_0$ test based on the normal approximation tends to give P values that are greater than the true P values.

Table 1. Monte Carlo properties of $\hat{\sigma}^2_{cn}$ and the normal approximation for T^*.

[F = Exponential (1), H = Exponential (λ)]

n	Ave. $\hat{\sigma}^2_{cn}$	s	$\hat{P}(T^* \le -z_{.10})$	$\hat{P}(T^* \le -z_{.05})$	$\hat{P}(T^* \le -z_{.01})$	$\hat{P}(T^* \ge z_{.10})$	$\hat{P}(T^* \ge z_{.05})$	$\hat{P}(T^* \ge z_{.01})$
				$\lambda = .1,\ t_0 = .6,\ \sigma^2_{c0} = .0459$				
100	.0434	.0017	.144	.076	.023	.068	.023	.003
150	.0443	.0014	.120	.080	.020	.094	.039	.005
200	.0447	.0012	.126	.071	.014	.078	.041	.005
				$\lambda = .1,\ t_0 = 1,\ \sigma^2_{c0} = .0372$				
100	.0343	.0043	.122	.071	.016	.098	.051	.009
150	.0354	.0036	.121	.065	.014	.103	.046	.010
200	.0357	.0029	.120	.065	.018	.111	.052	.010
				$\lambda = 1/3,\ t_0 = .6,\ \sigma^2_{c0} = .0532$				
100	.0498	.0028	.137	.075	.027	.068	.027	.005
150	.0511	.0022	.180	.099	.029	.062	.022	.000
200	.0514	.0019	.121	.065	.012	.076	.036	.006
				$\lambda = 1/3,\ t_0 = 1,\ \sigma^2_{c0} = .0478$				
100	.0425	.0062	.127	.072	.021	.100	.048	.004
150	.0447	.0051	.119	.066	.018	.095	.052	.009
200	.0454	.0046	.107	.063	.017	.101	.042	.005

3. EFFICIENCY LOSS DUE TO CENSORING

In this section we study the efficiency loss from censoring by comparing the efficacy of the test based on $T_n = T(\hat{G}_n)$ for the uncensored model with the efficacy of the test based on T_n^c for the randomly censored model. Also we present a monotonicity property of the efficiency as the amount of censoring increases.

Since T_n^c and T_n have the same asymptotic means, the value of 1-R, where

$$R \equiv e_{F,H}(T_n^c, T_n) = \sigma_0^2/\sigma_{c0}^2,$$

can be taken as a measure of the efficiency loss due to censoring (cf. Chen, Hollander, and Langberg, 1983a). Here σ_0^2 and σ_{c0}^2 are the null asymptotic variances of $n^{1/2}T_n$ and $n^{1/2}T_n^c$ respectively.

We consider the case where the censoring distribution is exponential with parameter λ and the life distribution is exponential with parameter 1. To satisfy Condition (A.2) we take $\lambda < 1$. Then we have

$$\sigma_0^2 = (1/12)e^{-t_0} + (1/12)e^{-2t_0} - (1/6)e^{-3t_0},$$

$$\sigma_{c0}^2 = \{1/4(3-\lambda)\} \cdot (e^{-(1-\lambda)t_0} + e^{-2t_0} - 2e^{-(3-\lambda)t_0}),$$

and

$$R = \{1-(\lambda/3)\}(1+e^{-t_0}-2e^{-2t_0})/(e^{\lambda t_0}+e^{-t_0}-2e^{-(2-\lambda)t_0}). \tag{3.1}$$

In Table 2, values of R given by (3.1) are presented for several choices of t_0 and $\lambda < 1$. The table shows that as λ decreases, the value of R increases to 1. Note that $\lambda = 0$ implies no censoring. The table also shows that as t_0 increases, the efficiency of T_n^c with respect to T_n decreases.

The next theorem shows that if H_2 is stochastically smaller than H_1 (i.e., there tend to be more censored observations with censoring distribution H_2 than with H_1), then the efficiency of T_n^c with respect to T_n under H_1 is greater than under H_2.

THEOREM 2. Assume that $H_1 \le H_2$, where H_1 and H_2 are censoring distributions. Then $R_1 \ge R_2$, where $R_1 = e_{F,H_1}(T_n^c, T_n)$ and $R_2 = e_{F,H_2}(T_n^c, T_n)$.

Proof. Since the numerator of $e_{F,H_i}(T_n^c, T_n)$ $i = 1,2$, does not depend on H, it suffices to show that $\sigma_{c0}^2(H_1) \le \sigma_{c0}^2(H_2)$ where $\sigma_{c0}^2(H)$ is given by (2.4). (Recall that $\bar{K} = \bar{F}\bar{H}$.) Since $H_1 \le H_2$ we have $[\bar{H}_1]^{-1} \le [\bar{H}_2]^{-1}$. Note that $\sigma_{c0}^2(H)$ is increasing in $[\bar{H}]^{-1}$. Thus, it follows that $\sigma_{c0}^2(H_1) \le \sigma_{c0}^2(H_2)$. Consequently, the desired result follows.

From Theorem 2, it immediately follows that the maximum value of R is equal to 1 and is achieved when $\bar{H}(x) \equiv 1$ for $x \ge 0$, that is, when there is no censoring.

Table 2. Efficiency of T_n^c with respect to T_n when $\bar{F}(x) = e^{-x}$, $x \ge 0$ and $\bar{H}(x) = e^{-\lambda x}$, $x \ge 0$, $\lambda > 0$

t_0 \ λ	1/2	1/4	1/8	1/20	1/60	1/100
.2	.9008	.9514	.9759	.9904	.9968	.9981
.6	.7265	.8583	.9280	.9709	.9903	.9942
1.0	.5823	.7711	.8804	.9509	.9835	.9901
1.4	.4652	.6907	.8338	.9307	.9765	.9858
2.0	.3325	.5843	.7673	.9003	.9657	.9793
3.0	.1931	.4442	.6686	.8520	.9482	.9686
4.0	.1146	.3411	.5854	.8077	.9314	.9583
5.0	.0688	.2639	.5146	.7670	.9154	.9484

4. AN EXAMPLE.

The data in Table 3 are an updated version of data analyzed by Koziol and Green (1976). The data correspond to 211 State IV prostate cancer patients treated with estrogen in a study by the Veterans Administration Cooperative Urological Research Group (1967). By the March 1977 closing date, 90 patients had died of prostate cancer, 105 had died of other diseases, and 16 were still alive. The latter 121 observations will be treated as censored observations (withdrawals).

Table 3. Survival times and withdrawal times in months for 211 patients (with number of ties given in parentheses)

Survival times: 0(3), 2, 3, 4, 6, 7(2), 8, 9(2), 11(3), 12(3), 15(2), 16(3), 17(2), 18, 19(2), 20, 21, 22(2), 23, 24, 25(2), 26(3), 27(2), 28(2), 29(2), 30, 31, 32(3), 33(2), 34, 35, 36, 37(2), 38, 40, 41(2), 42(2), 43, 45(3), 46, 47(2), 48(2), 51, 53(2), 54(2), 57, 60, 61, 62(2), 67, 69, 87, 97(2), 100, 145, 158.

Withdrawal times: 0(6), 1(5), 2(4), 3(3), 4, 6(5),7(5), 8, 9(2), 10, 11, 12(3), 13(3), 14(2), 15(2), 16, 17(2), 18(2), 19(3), 21, 23, 25, 27, 28, 31, 32, 34, 35, 37, 38(4), 39(2), 44(3), 46, 47, 48, 49, 50, 53(2), 55, 56, 59, 61, 62, 65, 66(2), 72(2), 74, 78, 79, 81, 89, 93, 99, 102, 104(2), 106, 109, 119(2), 125, 127, 129, 131, 133(2), 135, 136(2), 138, 141, 142, 143, 144, 148, 160, 164(3).

Koziol and Green (1976) stated that experience had suggested that had the patients not been treated with estrogen, their survival distribution F for deaths from cancer of the prostate could be taken to be exponential with a mean of 100 months. With this in mind, various authors have applied goodness-of-fit tests to the prostate cancer data. A recent reference is Csörgő and Horváth (1981) whose procedures indicate lack of support for the simple null hypothesis that F is exponential with mean 100. (Csörgő and Horváth also discuss other references relating to tests of this simple null hypothesis.)

Tests of the composite null hypothesis of exponentiality (with unspecified mean) were performed by Chen (1981), Chen, Hollander, and Langberg (1983a), and Chen, Hollander, and Langberg (1983b). These tests indicated lack of support for the hypothesis of exponential aging. Chen, Hollander, and Langberg (1983b) also plotted an empirical mean residual life function for the data of Table 3. Their plot tends to decrease up to around 25 months, then tends to increase up to about 70 months, and then decreases again.

The null hypothesis H_0 (1.3) of this paper is appropriate if one has *a priori* reasons to suspect that a patient after t_0 months would have stochastically the same residual lifelength as a new patient. However, we are not aware of any such *a priori* notions in this prostate cancer setting, and there is no natural value of t_0. Hence, our test is performed primarily for purposes of illustration. With the choice $t_0 = 60$, we find $T_{211}^c = .0116$, $\hat{\sigma}_{c211}^2 = .3354$, and $(211)^{1/2} T_{211}^c \hat{\sigma}_{c211}^{-1} = .290$, a value supporting H_0.

ACKNOWLEDGEMENTS

We are grateful to Frank Guess and Jim Sconing for checking the efficiency values of Table 2 and the value of our test statistic applied to the data of Table 3. We are grateful to Jim Sconing for performing the Monte Carlo study reported in Table 1. This research was supported by the United States Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR 82-K-0007 to the Florida State University.

REFERENCES

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship, Ann. Statist. 2, 437-453.

Chen, Y.Y., Hollander, M. and Langberg, N.A. (1982). Small-sample results for the Kaplan-Meier estimator, J. Am. Statist. Assoc. 77, 141-144.

Chen, Y.Y., Hollander, M. and Langberg, N.A. (1983a). Testing whether new is better than used with randomly censored data, Ann. Statist. 11, 267-274.

Chen, Y.Y., Hollander, M. and Langberg, N.A. (1983b). Tests for monotone mean residual life, using randomly censored data, Biometrics 39, 119-127.

Csörgő, S. and Horváth, L. (1981). On the Koziol-Green model for random censorship, Biometrika 68, 391-401.

Efron, B. (1967). The two-sample problem with censored data, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability Vol. IV, University of California Press, Berkeley, California, 831-852.

Gill, R.D. (1983). Large sample behaviour of the product-limit estimator on the whole line, Ann. Statist. 11, 49-58.

Hollander, M., Park, D.H. and Proschan, F. (1983). Testing whether new is better than used of a specified age (submitted).

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, J. Am. Statist. Assoc. 53, 457-481.

Koziol, J.A. and Green, S.B. (1976). A Crámer-von Mises statistic for randomly censored data, Biometrika 63, 465-474.

Langberg, N.A., Proschan, F. and Quinzi, A.T. (1981). Estimating dependent life lengths with applications to the theory of competing risks, Ann. Statist. 9, 152-167.

Marsaglia, G. and Tubilla, A. (1975). A note on the 'lack of memory' property of the exponential distribution, Ann. Prob. 3, 353-354.

Meier, P. (1975). Estimation of a distribution function from incomplete observations, in Perspectives in Probability and Statistics, J. Gani (ed.), 67-87. Sheffield: Applied Probability Trust.

Peterson, A.V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions, J. Am. Statist. Assoc. 72, 854-858.

Veterans Administration Cooperative Urological Research Group (1967). Treatment and survival of patients with cancer of the prostate, Surgery, Gynecology, and Obstetrics, 124, 1011-1017.

END

FILMED

9-84

DTIC